

FIGURE 1 A
NUCLEOTIDE SEQUENCE

>protease D-G (SEQ.ID.NO.:1)

CAACTTCACTTGTAGGGCTGTTTTAATCAAGCTGCCCAAAGTCCCCCAATCACTCCTGGA
ATACACAGAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAA
GGCCTGCCCTGCACTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTG
GCCAGCCAGGACCTGTGTGGGGAGGCCCTCCTGCTGCCTTGGGGTGACAATCTCAGCTCC
AGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGGATCCTGACAGTGATCAACCT
CTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGA
AAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGTGGTT
GTCCTCATCAAGGTGATTCTGGATAAATACTACTTCCCTCTGCGGGCAGCCTCTCCACTTC
ATCCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAG
CACTGTGTCAAGAGCTTCCCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGA
TCCCACTGCAGGTGCTGGACTCGGCCACAGGGAAGTGGTTCTCTGCCTGTTTCGACAAC
TTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGGGCTACAGCAGCAAACCCACT
TTCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAAC
AGCCAGGAGCTTCGCATGCGGAAGTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCC
CTGCACTGTCTTGCTGTGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGGAGGAG
GCCTCTGTGGATTCTTGGCCTTGGCAGGTACAGTACGACAAACAGCACGTCTGT
GGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCAGCCCACTGCTTCAGGAAACAT
ACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAGCTTCCCATCC
CTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAAAGACAATGAC
ATCGCCCTCATGAAGCTGCAGTTCCTCACTTCTCAGGCACAGTCAGGCCCATCTGT
CTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGC
TTTACGAAGCAGAATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTC
ATTGACAGCACACGGTGCAATGCAGACGATGCGTACCTGGGGGAAGTCACCGAGAAGATG
ATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCC
CTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGC
GGGGGCCCCGAGCACCCAGGGGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTAC
AATGTCTGGAAGGCTGAGCTGTAATGCTGCTGCCCTTTGCAGTGCTGGGAGCCGCTTCC
TTCCTGCCCTGCCACCTGGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCTTGG
GTACACCCCTCTGCCACAGCCTCAGCATTTCTTGGAGCAGCAAAGGGCCTCAATTCCTA
TAAGAGACCCTCGCAGCCCAGAGGCGCCAGAGGAAGTCAGCAGCCCTAGCTCGGCCACA
CTTGGTGCTCCCAGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCTCAGGGGTAT
TGCTAAGCCAAGAAGGAACTTTCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAG
CCCAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTT
CACCCATCCCCAAGCCTACTAGAGCAAGAAACAGTTGTAATATAAAATGCACTGCCTAC
TGTTGGTATGACTACCGTTACCTACTGTTGTCATTGTTATTACAGCTATGGCCACTATTA
TTAAAGAGCTGTGTAACATCA

0650745-063000

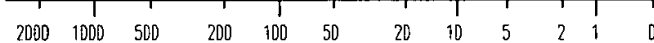
FIGURE 1 B
AA SEQUENCE

>protease D-G (SEQ.ID.NO.:2)

MDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVVLIKVILDKYYF
LCGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGN
WFSACFDNFTEALAEACRQMGYSSKPTFRAVEIGPDQDLVDVEITENSQELMRNSSGP
CLSGSLVSLHCLACGKSLKTPRVVGGEASVDSWPWQVSIQYDKQHVC GGSILDPHWVLT
AAHCFRKHTDVFNWKVRAGSDKLGSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTF
SGTVRPICLPFFDEELTPATPLWIIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAY
QGEVTEKMMCAGIPEGGVDTCCGDSGGPLMYQSDQWHVVGIVSWG YGCGGPSTPGVYTKV
SAYLNWIYNVWKAEL

000290:54220960

PHYLOGENETIC TREE



000250-54220560

FIGURE 3
TISSUE DISTRIBUTION

Protease D-G mRNA Tissue Distribution

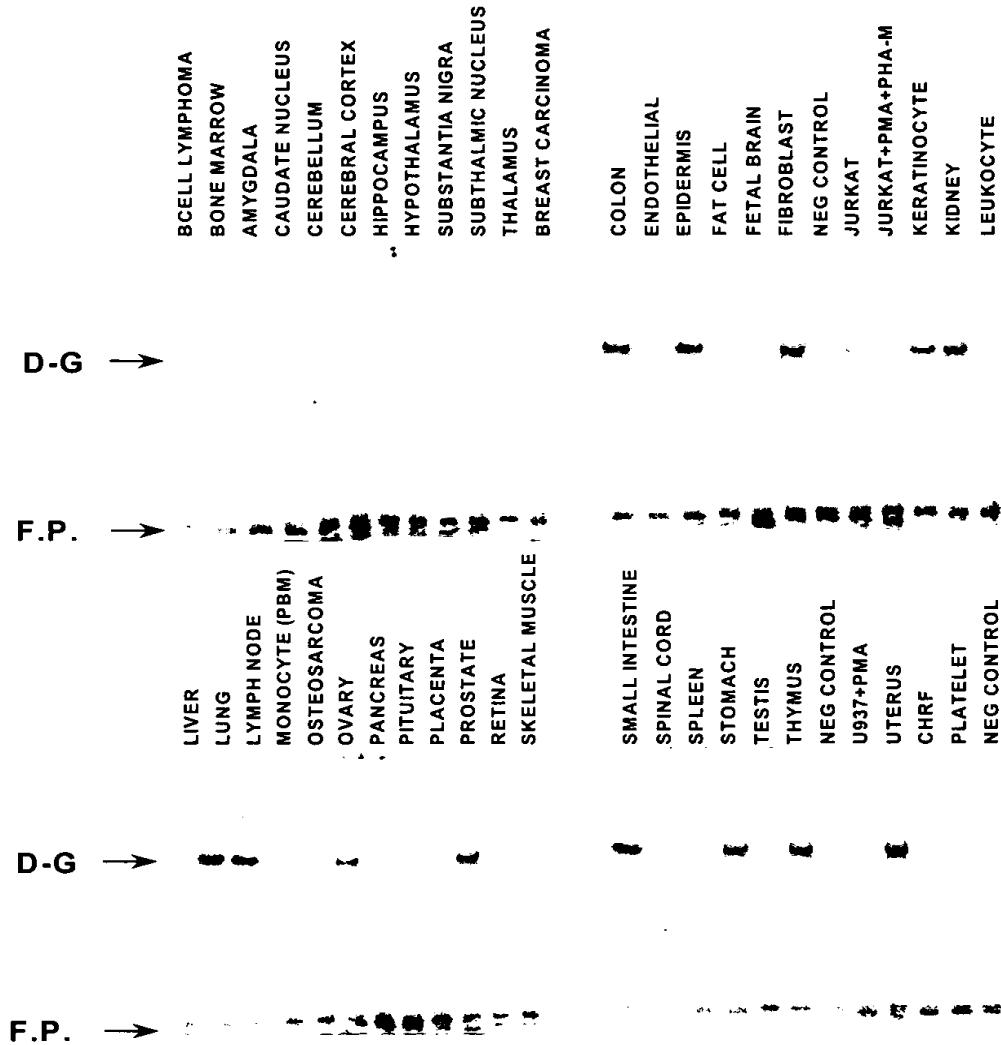


FIGURE 4 -A

CONSTRUCT NUCLEOTIDE SEQUENCE

>PFEK-D-G-HIS ERI-HCII (SEQ.ID.NO.:8)

GAATTCACCACCATGGACAGCAAAGGTTTCGTTCGACAGAAATCCCGCCTGCTCCTGCTGCTG
GTGGTGTCAAATCTACTCTTGTGCCAGGGTGTGGTCTCCGACTACAAGGACGACGACGAC
GTGGACGCGGCCGCTCTTGCTGCCCCCTTTGATGATGATGACAAGATCGTTGGGGGCTAT
GCTCTAGATGTGGATTCTTGCCCTTGGCAGGTCAGCATCCAGTACGACAAACAGCACGTC
TGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCAGCCCACTGCTTCAGGAAA
CATAACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGGCAGCTTCCCA
TCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAAAGACAAT
GACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATC
TGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGG
GGCTTTACGAAGCAGAATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCAG
GTCATTGACAGCACACGGTGCAATGCAGACGATGCGTACCTGGGGGAAGTCACCGAGAAG
ATGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGG
CCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGC
TGCGGGGGCCCGAGCACCCAGGGGTATACACCAAGGTCTCAGCCTATCTCAACTGGATC
TACAATGTCTGGAAGGCTGAGCTGTCTAGACATCACCATCACCATCACTAGCGGCCGCTT
CCCTTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGAC
AAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTG
CTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAGCTTGTCGAGAAGTACTAGAG
GATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACA
CCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAAC

FUSION PROTEIN

>PFEK-D-G-HIS (SEQ.ID.NO.:9)

MDSKGSSQKSRLLLL LVVSNLLLCQGVVSDYKDDDDVDAAALAAPFDDDDKIVGGYALDVDS
WPWQVSIQYDKQHVC GGSILDPHWVLTAAHCFRKHTDVFNWKVRAGSDKLGSFPSLAVAKII
IIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIIGWGFTKQNGGKM
SDILLQASVQVIDSTRCNADDA YQGEVTEKMMCAGIPEGGVDTCQGDSGGPLMYQSDQWHVV
GIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAELSRHHHHH

[illegible]

GAATTACCACCATGGACAGCAAGGTTGCTGCGAGAAATCGGCTGCTGCTGCTGGTGTGCTCAAAATCTACTCTGTGCGCCAGGGTGTGGTCTCCG
CTTAAGTGGTGGTACCTGCTGCTTTCCAAGCAGCGTCTTTAGGGCGGACGAGGACGACGACCACCACAGTTTAGATGAGAACACGGTCCACACCAGAGGC 100

M D S K G S S O K S R L L L L L V V S N L L L C Q G V V S
Prolactin Signal Sequence

Not I Xba I

ACTACAAGGACGACGACGACGCTGGACGCGGCCGCTCTTGCTGCCCCCTTTGATGATGATGACAAGATCGTTGGGGCTATGCTCTAGATGTGGATTCTTG
TGATGTTCTGCTGCTGCTGCTGCACTGCGCCGGCGAGAACGACGGGGAAACTACTACTGTTCTAGCAACCCCGATACGAGATCTACACCTAAGAAC 200

FLAG EK

D Y K D D D D V D A A A L A A P F D D D D K I V G G Y A L D V D S W

GCCTTGGCAGGTCAAGTCCAGTACGACAAACAGCACGCTGTGGAGGGAGCATCCTGGACCCCACTGGGTCTCACGGCAGCCCACTGCTTCAGGAAA
CGAACCGTCCAGTCTAGTCACTGCTGTTGTGCTGCGACACCTCCCTCGTAGGACCTGGGGGTGACCCAGGAGTCCGTCGGGTGACGAAGTCTTTT 300

P W Q V S I O Y D K Q H V C G G S I L D P H W V L T A A H C F R K

Protease D-G Catalytic Domain

CATACCGATGTGTTCAACTGGAAGTGCGGGCAGGCTCAGACAAATGGGCAGCTTCCCATCCTGGCTGTGGCCAAGATCATCATCAATTGAATTCAACC
GATGGCTACACAAGTTGACCTTCCACGCGCTCGAGTCTGTTGACCCGTCGAAGGTAGGGACCGACACCGGTTCTAGTAGTAGTAACCTAAGTTGG 400

H T D V F N W K V R A G S D K L G S F P S L A V A K I I I I E F N

Protease D-G Catalytic Domain

CCATGTACCCAAAGACAATGACATCGCCCTCATGAAGTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGCTGCCCTTCTTTGATGA
GGTACATGGGGTTTCTGTTACTGTAGCGGGAGTACTTCGACGCTCAAGGGTGAGTGAAGAGATCCGTGTCAGTCCGGGTAGACAGACGGGAAGAACTACT 500

P M Y P K D N D I A L M K L O D F P L T F S G T V R . P I C L P F F D E

Protease D-G Catalytic Domain

GGAGTCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCAG
CCTCGAGTGAGGTCCGTGGGGTGAGACCTAGTAACCTACCCGAAATGCTTCGCTTACCTCCCTTCTACAGACTGTATGACGACGTCGCCAGTCAGGTC 600

E L T P A T P L W I I G W G F T K Q N G G K M S D I L L O A S V O

Protease D-G Catalytic Domain

GTCATTGACAGCACACGGTGCAATGCAGACGATGCGTACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCT
CAGTAATGTCGTGTGCCACGTTACGCTGCTACGATGCTCCCTTCACTGGCTCTTCTACTACACAGTCCGTAGGGCCTTCCCCACACCTGTGGA 700

V I D S T R C N A D D A Y O G E V T E K M H M C A G I P E G G V D T

Protease D-G Catalytic Domain

GCCAGGGTGACAGTGGTGGGCCCTGATGTACCAATCTGACCAAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGGGGGCCCGAGCACCCC
CGGTCCCACTGTCAACACCCGGGGACTACATGGTTAGACTGGTCACCGTACACCACCCGTAGCAATCGACCCGATACCGACGCCCCGGGGCTCGTGGGG 800

C O G D S G G P L M Y O S D O W H V V G I V S W G Y G C G G P S T P

Protease D-G Catalytic Domain

AGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTGTCTAGACATCACCATCACCATCACTAGCGGCCGCTT
TCCTCATATGTGGTCCAGAGTCGGATAGAGTTGACCTAGATGTTACAGACCTTCCGACTCGACAGATCTGTAGTGGTAGTGGTAGTGATCGCCGGCGAA 900

G V Y T K V S A Y L N W I Y N V W K A E L S R H H H H H H

Protease D-G Catalytic Domain 6XHis-TAG

CCCTTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTG
GGGAAATCACTCCCAATTACGAAGCTCGTCTGTACTATTCTATGTAACCTACTCAAACCTGTTTGGTGTGTGATCTTACGTCACCTTTTTTTACGAAATAAAC 1000

TGAAATTTGTGATGCTATTGCTTTATTTGTAACCATATAAGCTGCAATAAACAAGTTAGCTTGTGCGAGAAGTACTAGAGGATCATAATCAGCCATACCA
ACTTTAAACACTACGATAACGAAATAAACATTGGTAATATTCGACGTTATTGTGTTCAATCGAACAGCTCTTCATGATCTCCTAGTATTAGTCGGTATGGT 1100

CATTTGTAGAGGTTTTACTTGCTTTAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACT 1189

GTAACATCTCCAAATGAACGAAATTTTTGGAGGGTGTGGAGGGGGACTTGGACTTTGTATTTTACTTACGTTAACAACAACAATTTG

FIGURE 5
PAGE - WESTERN BLOT

Protease D-G

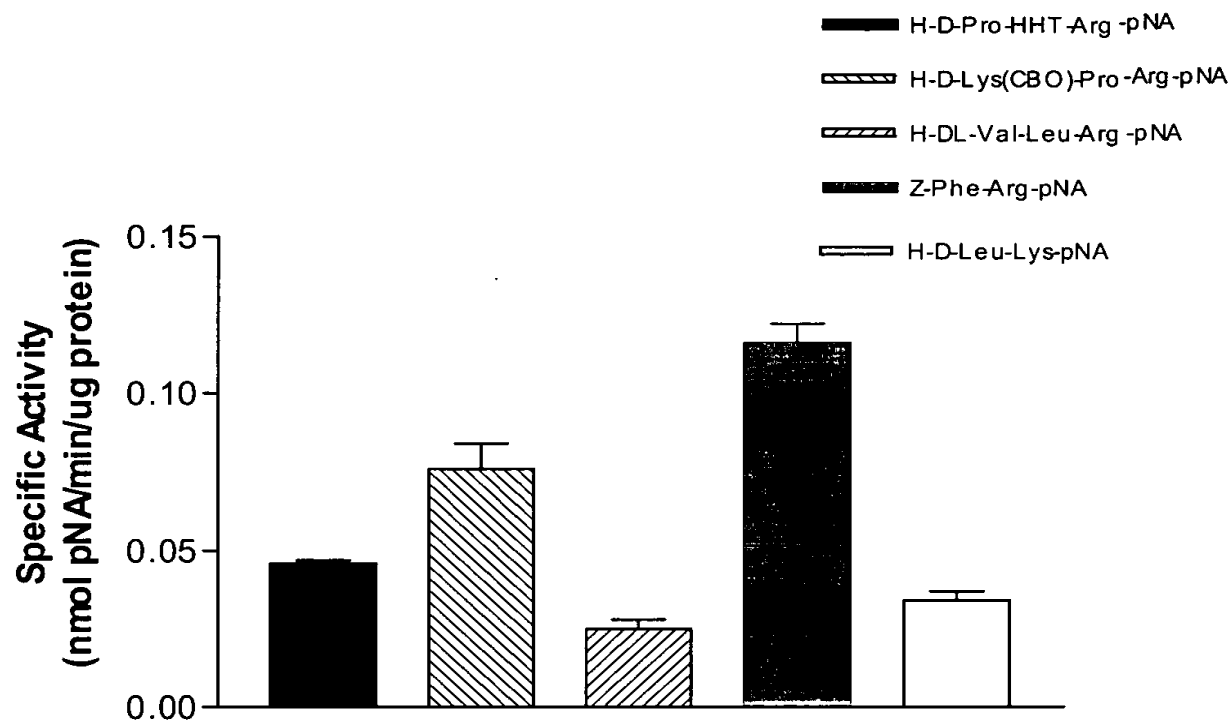
1 2 3 4 5
- + - + EK

111.0
73.0
47.5
33.9
28.8

De la

FIGURE 6

Chromogenic Activity of Protease D-G



H = free amine
 D = D-isomer
 CBO = carbobenzoxy
 HHT = hexahydrotyrosyl
 Z = benzyloxycarbonyl
 pNA = p-nitroanilide
 DL = D- and L- isomers